Vision Research 50 (2010) 1559–1570

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The effective fusional range for words in a natural viewing situation

Hazel I. Blythe [a,*], Simon P. Liversedge [a], John M. Findlay [b]

[a] School of Psychology, Shackleton Building, University of Southampton, Highfield, Southampton, SO17 1BJ, UK
[b] Psychology Department, Durham University, Science Laboratories, South Road, Durham, DH1 3LE, UK

ARTICLE INFO

Article history:
Received 1 September 2009
Received in revised form 13 May 2010

Keywords:
Retinal disparity
Binocular coordination
Reading

ABSTRACT

We investigated the effective fusional range for written stimuli in children and adults in a natural viewing situation. We recorded binocular eye movements in children and adults during processing of stereoscopically presented words in a lexical decision task. The effect of disparity magnitude on ease of fusion caused decreased response accuracy, increasing numbers of fixations and increased trial viewing times when retinal disparity exceeded one character space. The data suggest that retinal inputs of a word that are disparate by up to one character (0.37°) fall within the effective fusional range such that lexical decisions are not impaired.

1. Introduction

The fusion of disparate retinal inputs from each eye is a fundamental process underlying human visual processing in order that a single, unified visual percept is experienced (Liversedge, Rayner, White, Findlay, & McSorley, 2006). Furthermore, humans sample their visual environment by orienting their eyes, making saccades (ballistic, coordinated movements of both eyes) and fixations (brief periods when the eyes are relatively still during which visual information is taken in). At a neurophysiological level, a complex system is responsible for the coordination of the eyes (Zee, Fitzgibbon, & Optican, 1992). The same oculomotor neurons control both conjugate and disconjugate eye movements (Cova & Galiana, 1995; Sylvestre, Choi, & Cullen, 2003); however, specialised neural systems drive disconjugate convergent or divergent adjustments of the vergence angle between the eyes (Judge & Cumming, 1986). Electrophysiological cellular recordings in monkeys and cats have shown that premotor neurons (termed vergence tonic cells) in regions including the primary visual cortex, middle temporal area, and medial superior temporal area, project to the oculomotoneurons and control disconjugate eye movements, coding specific vergence angles (Barlow, Blakemore, & Pettigrew, 1967; DeAnglis & Newsome, 1999; Leigh & Zee, 1999; Pettigrew, Nikara, & Bishop, 1968).

A significant body of research has now been conducted to examine binocular coordination in humans (e.g., Bains, Crawford, Cadera, & Vilis, 1992; Collewijn, Erkelens, & Steinman, 1988, 1995; Erkelens & Sloot, 1995; Kloke & Jaschinski, 2006; Zee et al., 1992; see Kirkby, Webster, Blythe, and Liversedge (2008) for a review), and a number of basic characteristics have been established. First, disconjugacy between the two eyes originates during saccades. Second, the magnitude of the disconjugacy (within limits) is related to the amplitude of the saccade, with greater disconjugacy associated with larger saccade amplitudes. Third, at fixation onset there is usually residual disparity between the location of the point of fixation of each eye that results from saccadic disconjugacy. Fourth, this fixation disparity is reduced through small, comparatively slow, vergence movements that persist throughout the duration of a fixation. Fifth, and finally, the vergence movements do not entirely eliminate fixation disparity – disparity persists such that on a significant proportion of fixations the lines of sight are not aligned at subsequent saccade onset.

Such research underpins current understanding of non-diplopic vision as a consequence of binocular coordination during sequential fixations and saccades (Collewijn et al., 1988; Enright, 1984; Kenyon, Ciuffreda, & Stark, 1980; King & Zhou, 2000). Eye movements serve to control the magnitude of fixation disparity in order to prevent the experience of diplopia (double vision). When light from an object falls on slightly different locations on the two retinas, this difference is referred to as retinal disparity. When retinal disparity occurs, we do not perceive two separate objects – the visual system is, more often than not, able to fuse the two retinal images so that a single object is perceived. There are, however, limits to the magnitude of retinal disparity that can be fused. Central to non-diplopic vision is the notion of Panum's fusional area – the measured magnitude of retinal disparity at different retinal eccentricities that can be successfully fused.

Early investigations that attempted to measure Panum's area for simple visual stimuli suggested that for central vision the limit

* Corresponding author. Fax: +44 23 8059 4597.
E-mail address: hib@soton.ac.uk (H.I. Blythe).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.05.011

for fusion lies between 0.1° and 0.2° (Ames, Ogle, & Gliddon, 1932; Ogle, 1952). Other experiments have also attempted to quantify Panum's fusional area (Fender & Julesz, 1967; Heckmann & Schor, 1989; Lee & Dobbins, 2006; Palmer, 1961; Richards & Foley, 1971; Schor & Tyler, 1981; Schor, Wood, & Ogawa, 1984). However, all these studies have either prevented participants making useful eye movements in relation to retinal disparity (in some cases through retinal stabilisation techniques), or have failed to record eye movements, thereby precluding the possibility of an accurate index of retinal disparity.

Furthermore, previous studies have often employed the method whereby images are presented uniquely to each eye, initially aligned so that they are fused, and are gradually drawn apart until fusion is broken. In these circumstances, retinal disparity is produced by moving the stimuli presented separately to each eye in relation to each other. Subsequently, when the diplopic stimuli are gradually moved towards each other again until fusion is regained, this typically produces a smaller estimate of Panum's area than when the two images are drawn apart. This suggests that fusion can be maintained to greater disparities than can more naturally be achieved in response to disparities. Thus, the method of stimulus presentation can affect measurements of Panum's fusional area.

Our particular concern is with the process of reading. Reading is a remarkable skill that is critical to successful function in today's highly literate and technological society. It involves complex visuo-cognitive processing, and a non-diplopic visual representation of the words in a sentence is a necessary pre-requisite for such processing to occur. During reading, the stimulus is stationary and fixed, and retinal disparity occurs due to differences in the fixation positions of the two eyes – fixation disparity. Studies have shown that fixation disparities during reading are often as large as 0.5° (Blythe et al., 2006; Liversedge, White, Findlay, & Rayner, 2006; Nuthmann & Kliegl, 2009), which is considerably larger than the fusional limits for point targets. Furthermore, disparity changes on a moment-to-moment basis from saccade and fixation to saccade and fixation during reading. Importantly, in this situation, the fixed stable vergence state that is conventionally required for the measurement of Panum's fusional area (Howard, 2001) does not occur.

In the present study, we investigated the effective fusional range for written linguistic stimuli (words and nonwords), showing the range of retinal disparities that can be consistently and reliably fused (i.e., perceived as non-diplopic), thereby allowing the participant to successfully identify a word in a natural viewing situation.[1] The precise nature of the visual stimulus can have a profound effect on the magnitude of disparity that can be fused (Burt & Julesz, 1980; Fender & Julesz, 1967; Heckmann & Schor, 1989; Schor & Tyler, 1981; Schor et al., 1984). Linguistic stimuli are visually complex, containing both high and low spatial frequencies and constituent letters comprised of a range of oriented visual features.

The examination of this issue is particularly important with respect to beginning readers since marked changes in binocular coordination occur through childhood that have been speculatively attributed to the development of non-specified neural or muscular systems (Blythe et al., 2006; Yang, Bucci, & Kapoula, 2002; Yang & Kapoula, 2003). Disparity magnitudes during fixations in reading are greater for children than adults (on average, 0.3° in children compared to 0.2° in adults) (Blythe et al., 2006; Kirkby et al., 2008). Furthermore, children have been reported to make a greater proportion of crossed (compared to uncrossed) fixations than adults (Blythe et al., 2006; Kirkby et al., 2008). Thus, the direction of retinal disparity that is most often experienced changes with age, for both reading and non-reading tasks (Fioravanti, Inchingolo, Pensiero, & Spanios, 1995). It seems entirely plausible, therefore, that such developmental changes in binocular coordination might be associated with developmental changes in binocular fusion, an issue that has not previously been investigated.

In this study, we tracked binocular eye movements whilst presenting visual linguistic stimuli (words and nonwords) stereoscopically. This combination of taking high-quality binocular eye movement recordings and using a stereoscopic stimulus presentation allowed us to assess how the process of binocular fusion during reading changes with development. Note, also, that this type of viewing situation allows us to observe and investigate the influence of dynamic aspects of vergence control continuously during both fixations and saccades, thereby allowing us to understand the consequences of those vergence movements in relation to successful fusion.

In each trial of our experiment, the participant first looked at a fixation cross on the left side of the screen. This disappeared after one second, and then a target word (or nonword) appeared in the centre of the screen. We manipulated an imposed (stereoscopic) horizontal disparity within these stimuli of either zero, one, or two character spaces (0°, 0.37°, 0.74°, respectively). That is to say, our stereoscopic presentation technique allowed us to offset the word or nonword that was presented separately to each eye by a relative amount of zero, one, or two characters. Participants were required to fixate the letter string and then decide whether the stimulus was a word or a nonword. On half of the trials, one of the two central letters of the word was substituted to form a nonword. To successfully detect these single-letter misspellings, fusion of the two retinal inputs was necessary.

The disparity between the two eyes' positions was calculated (off-line) on a fixation-by-fixation basis, and we will refer to this as fixation disparity. Critically, by combining the two sources of disparity – manipulated stereoscopic disparity (on the screen) and the measured fixation disparity (naturally occurring between the eyes) – on a fixation-by-fixation basis, we were able to calculate a measure of retinal disparity for our statistical analyses (see Section 2.5). This constitutes a veridical measure of the difference in position of the stimuli falling on the two retinas and is, therefore, the most accurate index of disparity in the signal provided by the two eyes to the brain.

In this way, we measured response accuracy and various eye movement parameters across a range of crossed and uncrossed retinal disparities for both adults and children to assess the extent to which disparity impacted on the process of fusion. Thus, this experimental arrangement allowed the investigation of the effective fusional range for linguistic stimuli in a natural viewing situation.

We made two clear predictions for the present study. First, we predicted that both adults and children would easily be able to fuse one character space of retinal disparity in either direction. Previous work has shown that disparity during fixations in reading is often up to 0.42° in adults and 0.53° in children (based on mean values plus one standard deviation) – greater than one character space in the present study (Blythe et al., 2006; Kirkby et al., 2008; Liversedge, White, et al., 2006). Given that readers do not typically experience diplopia when reading, one character space of retinal disparity should result in fusion for both adults and children. Sec-

[1] To have measured Panum's fusional area, it can be argued that it is necessary to control the vergence state. To do this, ordinarily an image is retinally stabilised and then the two stimuli presented separately to each eye are manipulated in relation to each other. In the present experiment, we examined a dynamic viewing situation, in which we allowed natural eye movements and we calculated retinal disparity on a fixation by fixation basis. Given this experimental situation, we are somewhat cautious with respect to our claims in relation to measuring Panum's fusional area for words (as conventionally defined). We, therefore, consider our study throughout this manuscript as being an investigation of the effective fusional range for written linguistic stimuli. We are grateful to Wolfgang Jaschinski and Erich Graf for very helpful discussions in relation to this matter.

ond, we predicted that children may able to successfully fuse a greater range of retinal disparities than adults. This second prediction was made on the basis of two prior experimental results – (1) one study has found that children could perceive depth from large-disparity random-dot stereograms more quickly than adults (Dowd, Clifton, Anderson, & Eichelman, 1980), and (2) children typically experience larger fixation disparities than adults when reading (Blythe et al., 2006).

2. Method

2.1. Participants

Ten adults aged 18–21-years, and 10 children aged 7–11-years took part in the experiment. Adult participants were all undergraduate student volunteers at Durham University, and children were all volunteers recruited from local schools.

2.2. Apparatus

Stereoscopic presentation of the target words was achieved through use of CRS FE1 shutter goggles that block visual input to each eye alternately every 8 ms (corresponding to a 120 Hz refresh rate). These were interfaced with a Pentium 4 computer and a Philips 21B582BH 21″ monitor on which the target words were displayed. The white letters had a luminance of 14.2 cd/m^2 while the black background had a luminance of .06 cd/m^2. The monitor had a P22 phosphor with decay rate to 0 in less than 1 ms. Binocular eye movement recordings were taken with two Fourward Technologies Dual Purkinje Image eye trackers. The position of both eyes was recorded every millisecond. The display monitor was set at a viewing distance of 100 cm. All words were presented in block capitals, in Courier New size 18 font. At the specified viewing distance, one character space subtended 0.37° (22.2 arc min) of visual angle. All participants bit on a wax dental mould and used forehead rests during the experiment, to eliminate head movements.

2.3. Materials and design

All participants viewed 100 trials, each trial consisting of a single 6-letter word. Ten counterbalanced files were created from the set of 100 words, so that every word appeared in each of the five experimental conditions, both correctly and with a letter substitution (so forming a nonword). The misspellings were always a single-letter substitution, of either the third or the fourth letter of the word (i.e. the misspelling was always in the centre of the word), and created an obvious nonword (for example, changing MOTHER to MOTKER) in order to be as sure as possible that the younger children would detect the nonwords easily if they were able to see them clearly. The words were controlled for age of acquisition such that the latest acquired word would typically be known by children aged 7-years (MRC Database, Coltheart, 1981).

The five stereoscopic conditions correspond to the imposed horizontal displacement between the images sent to the two eyes (stereoscopic disparity). Words could be presented in the following five ways: (1) uncrossed by two character spaces; (2) uncrossed by one character space; (3) aligned, where the two images were in the same location on the display; (4) crossed by one character space; (5) crossed by two character spaces. In conditions 1 and 2, the image presented to the left eye was shifted to the left on the display relative to the image presented to the right eye. In conditions 4 and 5, the image presented to the left eye was shifted to the right of the image presented to the right eye.

2.4. Procedure

All participants were given both written and verbal instructions upon arrival. Participants were instructed to look at the fixation cross which appeared on the left of the screen for 1 s at the beginning of each trial, before looking at the word which was presented in the centre of the screen simultaneously with the offset of the fixation cross. The distance between the fixation cross and the left edge of the word was 1.3°. Participants were instructed to decide as quickly and accurately as possible for each trial whether the word was or was not spelled correctly and respond accordingly using a button box. The sequence of events in each trial is illustrated in Fig. 1, Panel b.

The left and right eye trackers were calibrated for each eye monocularly in turn (i.e., during calibration of the right eye, the left eye was occluded and vice versa). Once both eyes had been calibrated accurately, the practise and experimental sentences were then presented. All participants had five practise trials in order to make sure they were fully familiar and comfortable with the procedure before the experimental trials began. Following every four trials, the calibration was checked for accuracy, and the eye trackers were recalibrated if necessary. All participants were given a break half way through the experiment, and additional breaks were given as often as required. The entire experiment lasted approximately 40 min for children, due to their need for frequent breaks, and 20 min for adults.

2.5. Analyses

Custom-designed software was used for the data analyses. Fixations were manually identified in order to avoid contamination by dynamic overshoots (Deubel & Bridgeman, 1995; Liversedge, White, et al., 2006). Crossed disparities, where the letter falling on the fovea of the left eye was further to the right within the printed word than the letter falling on the fovea of the right eye, were denoted by positive values. For the converse situation, uncrossed disparities, values were denoted by negative values. For every fixation, retinal disparity was calculated by subtracting stereoscopic disparity from fixation disparity, thus giving a measure of disparity that encompassed both sources (see Panel a in Fig. 1 for the full set of retinal disparities that results from possible fixation and stereoscopic disparity combinations).

Convergence movements were deemed to have occurred if the eyes were more converged at the end of a specified period (either a fixation or a saccade) than they were at the beginning. Similarly, a divergent movement was deemed to have been made if the eyes were more diverged at the end of a specified period than they were at the beginning. We did not apply a minimum magnitude of movement as a criterion for defining vergence movements in order to maximise the statistical power of these analyses; thus, fewer than 1% of fixations and saccades were excluded from the analyses of vergence movements.

For all significant main effects of disparity, post hoc paired-samples *t*-tests were conducted. Based on the a priori expectation that participants would find it increasingly difficult to fuse both crossed and uncrossed disparities of increased magnitude, the four levels of disparity were each compared individually to the aligned condition. For interactions between participant group and disparity, the same four *t*-tests were conducted for each group separately.

3. Results

Below we report a variety of analyses, some of which are descriptive, some based on stereoscopic disparity (the stimulus manipulation) and others based on retinal disparity (the combina-

a

		Fixation disparity					
		FU3	FU2	FU1	FA	FC1	FC2
Stereoscopic Disparity	SU2	TU1	TA	TC1	TC2	TC3	TC4
	SU1	TU2	TU1	TA	TC1	TC2	TC3
	SA	TU3	TU2	TU1	TA	TC1	TC2
	SC1	TU4	TU3	TU2	TU1	TA	TC1
	SC2	TU5	TU4	TU3	TU2	TU1	TA

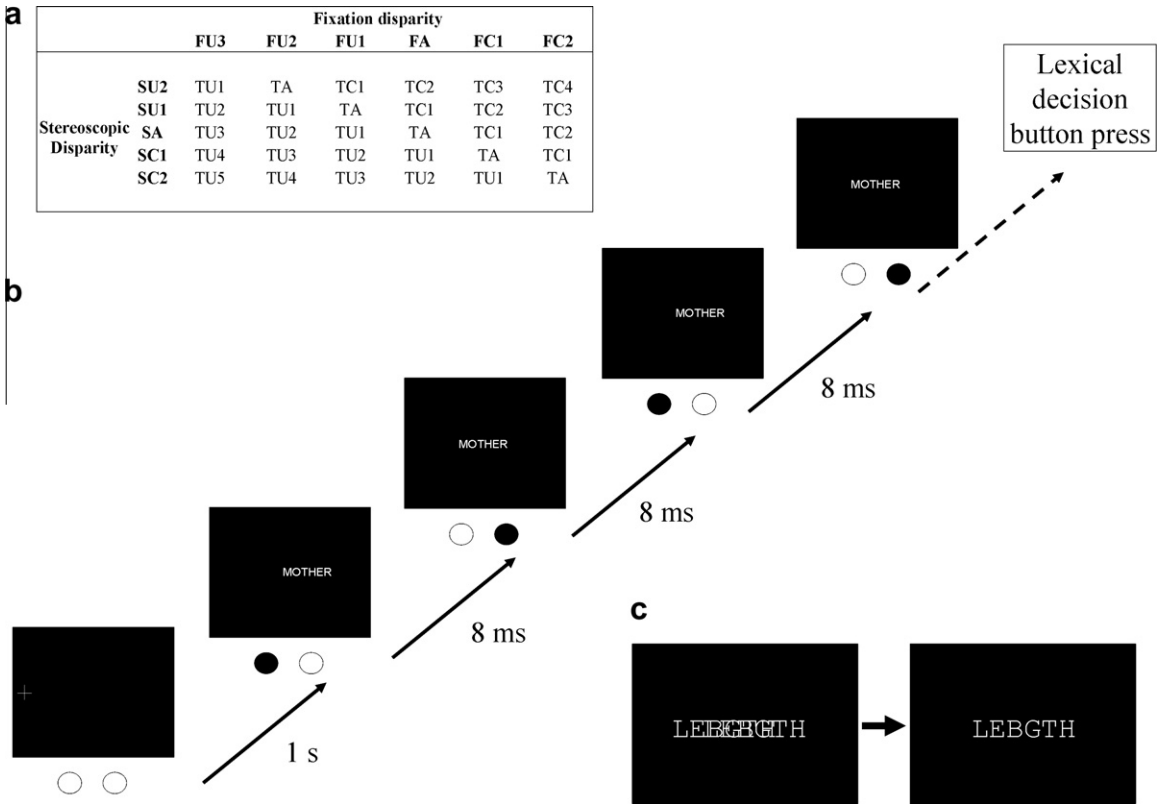

Fig. 1. Panel a shows how retinal disparity was calculated on a fixation-by-fixation basis, by combining fixation disparity (F) and stereoscopic disparity (S). The letter T refers to total retinal disparity, and crossed disparities are demarked by a (C), uncrossed disparities by a (U) and aligned inputs by an (A). The digit after the letter U or C denotes the magnitude of the disparity in character spaces. The formula for combining fixation disparity and stereoscopic disparity is given in Section 2.5. Panel b shows a representation of the sequence of events in each trial (note that this is not to scale). The circles, white and black, represent the shutter goggles as open and closed, respectively. The shutter goggles opened and closed every 8 ms, and this was synchronised with the alternating presentation of two bitmaps (either with or without a horizontal displacement depending on the presentation condition) on the display screen. The example here demonstrates an uncrossed stimulus, where the word presented to the left eye is shifted left by one character on the screen, relative to the position of the word presented to the right eye (note the differing distance between the end of the word and the edge of the screen for the two bitmaps). The trial ended when the participant pressed a button to indicate whether the stimulus was a word or a nonword. Panel c shows an example stimulus with a two character space disparity. The image on the left shows how the non-fused stimulus appeared, and the image on the right shows how the same stimulus appeared after fusion.

tion of stereoscopic and fixation disparity). We carried out these different types of analyses in order to address specific theoretical questions.

One of our primary objectives in this experiment was to evaluate adults' and children's effective fusional range for linguistic stimuli. For this reason, we first report response accuracy as a function of retinal disparity, as this is our primary index of successful fusion (Section 3.1).

As shown in Panel a of Fig. 1, we combined the manipulated stereoscopic disparity and the measured fixation disparity on a fixation-by-fixation basis in order to calculate our measure of total retinal disparity. Throughout the remainder of the paper we will consistently use specific labels to refer to the direction and magnitude of disparity. We characterised total retinal disparity (T) in terms of character spaces. Crossed disparities are denoted by a (C), uncrossed disparities by a (U) and aligned inputs by an (A). The digit after the letter U or C denotes the magnitude of the disparity in character spaces. Total retinal disparities formed through the combination of binocular fixation differences and stereoscopic display manipulations covered a range between TU4 and TC3 character spaces. However, there were very few data points at the more extreme disparities and, therefore, the analyses reported here are restricted to the central range (TU2 to TC2 character spaces; ±0.74°). Only 1% of adult fixations and 2% of child fixations were excluded due to falling outside this range, and the removal of these data points had a negligible influence on the statistical analyses and mean reported values.

In Sections 3.2 and 3.3, we examined whether participants were sensitive to stereoscopic disparity cues in the parafovea. At the beginning of each trial the participants fixated the cross on the left of the screen, and then planned and executed a saccade onto a parafoveal stereoscopic stimulus. We examined the vergence change during these initial saccades onto and initial fixations on the stimuli (Section 3.2), and landing positions on the stimuli (Section 3.3), as a function of stereoscopic disparity in order to examine whether participants were sensitive to parafoveal disparity cues.

In Sections 3.4 and 3.5 we report the time course of fusion and vergence movements during fixations on the stimulus in order to examine differences between adults and children. Here, it is more appropriate to consider retinal disparity (encompassing both stereoscopic and fixation disparity), rather than stimulus disparity as it is this that must be reduced in order for successful fusion to occur. Our measure of retinal disparity represents the alignment of the two eyes in relation to the stimulus in depth. For example, a stimulus presented with an uncrossed stereoscopic disparity of two character spaces would appear slightly behind the depth of the screen. If the participant had an uncrossed fixation disparity of two characters (as measured at the depth of the screen) then their eyes would actually be aligned at the apparent depth of the stimulus and, hence, they would have a retinal disparity of zero (again, see Panel a in Fig. 1 for all combinations). Thus, in Sections 3.4 and 3.5, where we consider differences in oculomotor behaviour between adults and children during fixations on the stimuli, we report retinal disparity measures.

In order to obtain a sense of participants' subjective experience during the experiment, after testing we asked each of them to provide a verbal report of how they found the task. When participants first fixated the letter strings they reported that it was clear that they were linguistic in nature. However, they reported that it was not always possible to see the constituent letters clearly or to read the word immediately, given that these stimuli literally appeared as double images with up to four overlapping letters. Participants reported that upon fixating the word, after a few moments it became entirely clear (albeit appearing either slightly in front of, or slightly behind the screen), and at this point they were able to make their response.

3.1. Response accuracy

Response accuracy was our behavioural index of whether or not participants had successfully fused the stereoscopic stimuli. The response accuracy data reflect the lexical decision that participants made under the different experimental conditions at the end of the final fixation of each trial (see Fig. 2). While participants often made multiple fixations during each trial (see Section 3.2) in which vergence movements occurred (see Section 3.5), the end of the final fixation on each trial corresponded to the moment at which the participant pressed the button to indicate their decision. Thus, disparity at the end of the final fixation of each trial is a measure of the remaining vergence error that has not been compensated for by eye movements during the course of the trial, and describes the effective fusional range.

Stimuli in the aligned stereoscopic condition (SA, see Panel a of Fig. 1) were presented with zero stereoscopic disparity (irrespective of fixation disparity) and were therefore representative of a word stimulus presented normally. These data therefore provide a baseline against which to make comparisons. Response accuracy in this condition was 99% for the adults and 89% for the children. These values represent the ability of our participants to detect single-letter misspellings within words presented normally, and show that, as would be expected, children performed slightly worse than adults on this task.

Note, however, our primary objective in this study was to investigate the effective fusional range for words in children and adults. Thus, we also carried out formal analyses of response accuracy as a function of retinal disparity (the disparity that must be fused in order for the participant to complete the lexical decision task accurately).

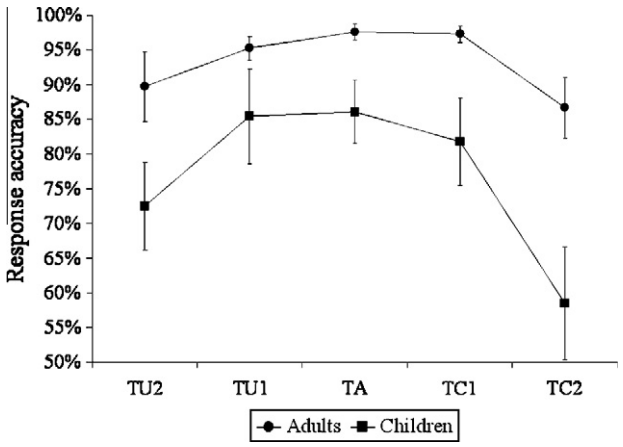

Fig. 2. Response accuracy across different values of retinal disparity at the end of the final fixation of each trial. The letter U denotes an uncrossed disparity; the letter C denotes a crossed disparity; A refers to aligned retinal inputs. The numbers denote the magnitude of the disparity in character spaces. Bars show the standard error for each participant group in each condition.

Clearly, response accuracy was poorest when the residual retinal disparity, at the end of the final fixation of each trial, was at its greatest (two character spaces).[2] This is particularly striking in the children's data, where the mean response accuracy decreased between disparities of one and two character spaces by 13% for uncrossed disparities, and 23% for crossed disparities.

There was a significant main effect of retinal disparity on response accuracy. When greater magnitudes of retinal disparity were experienced at the end of the final fixation of the trial, response accuracy was poorer compared to the case of aligned retinal inputs ($F(4, 72) = 2.53$, $p = .05$). As described above, response accuracy was poorest with two character spaces' retinal disparity in both the uncrossed direction ($t(19) = 2.22$, $p = .04$) and crossed direction ($t(19) = 2.40$, $p = .03$). These data show that when two character spaces of retinal disparity remained at the end of the final fixation of the trial, response accuracy decreased significantly. This suggests that participants had failed to successfully fuse the stimuli most often in these conditions.

In contrast, the differences between the one character space conditions (both crossed and uncrossed) and the aligned condition was much smaller and did not approach significance (both ts < 2, both ps > 0.2). Thus, when one character space of retinal disparity remained at the end of the final fixation of the trial, response accuracy was not reliably different from when the two retinal inputs were aligned to within one character space of each other. This indicates that, in these cases, participants were able to fuse the stimuli and so accurately complete the task.

There was an overall difference between adults and children – as might be expected, children were generally poorer than adults at detecting misspellings ($F(4, 18) = 6.66$, $p = .02$). This is very likely due to children being less skilled than adults in processing visual linguistic stimuli and therefore more error prone in relation to making lexical decisions. Note, however, that the overall difference in performance is not the issue of primary importance in relation to our central theoretical claims. Instead, it is the relative decrement in response accuracy compared to optimal performance with increased retinal disparity for adults and for children that is critical in the examination of fusion. However, the interaction between participant group and retinal disparity did not approach significance ($F < 1$). That is to say, response accuracy decreased in relation to increased retinal disparity in a comparable manner for both adults and children. In contrast to our predictions, therefore, children's ability to fuse two character spaces of retinal disparity was equivalent to that of adults. For both participant groups, there was a significant decrease in response accuracy on trials where there was a residual retinal disparity of two character spaces (both crossed and uncrossed) at the end of the trial.

It should be noted that there were relatively few trials where the participant was still experiencing two character spaces of retinal disparity at the end of the final fixation. At this moment, participants had made magnitude and direction-appropriate vergence movements during the majority of fixations in order to fuse the stimulus (see Section 3.5). These vergence movements served to

[2] Due to our use of high contrast stimuli (white text on a black background) and shutter goggles in the stereoscopic presentations, it is possible that cross-talk between the shutter goggles may have resulted in very faint traces of one eye's unique stimulus being visible to the other eye. To minimise any influence of such traces, we used FE1 shutter goggles, which have less cross-talk than many other types, and we also set the monitor to have a low contrast (see the luminance values for the white text and black background in Section 2.2). We do not believe that any of our effects were influenced by cross-talk because stereoscopic disparity had the predicted, systematically differential influence across stereoscopic conditions whilst any cross talk would be present in all stereoscopic conditions. To this extent, while we accept that cross-talk can be an issue of concern in relation to the use of shutter goggles for stereoscopic presentations, we do not believe that it was a determinant of the effects that we report in this paper.

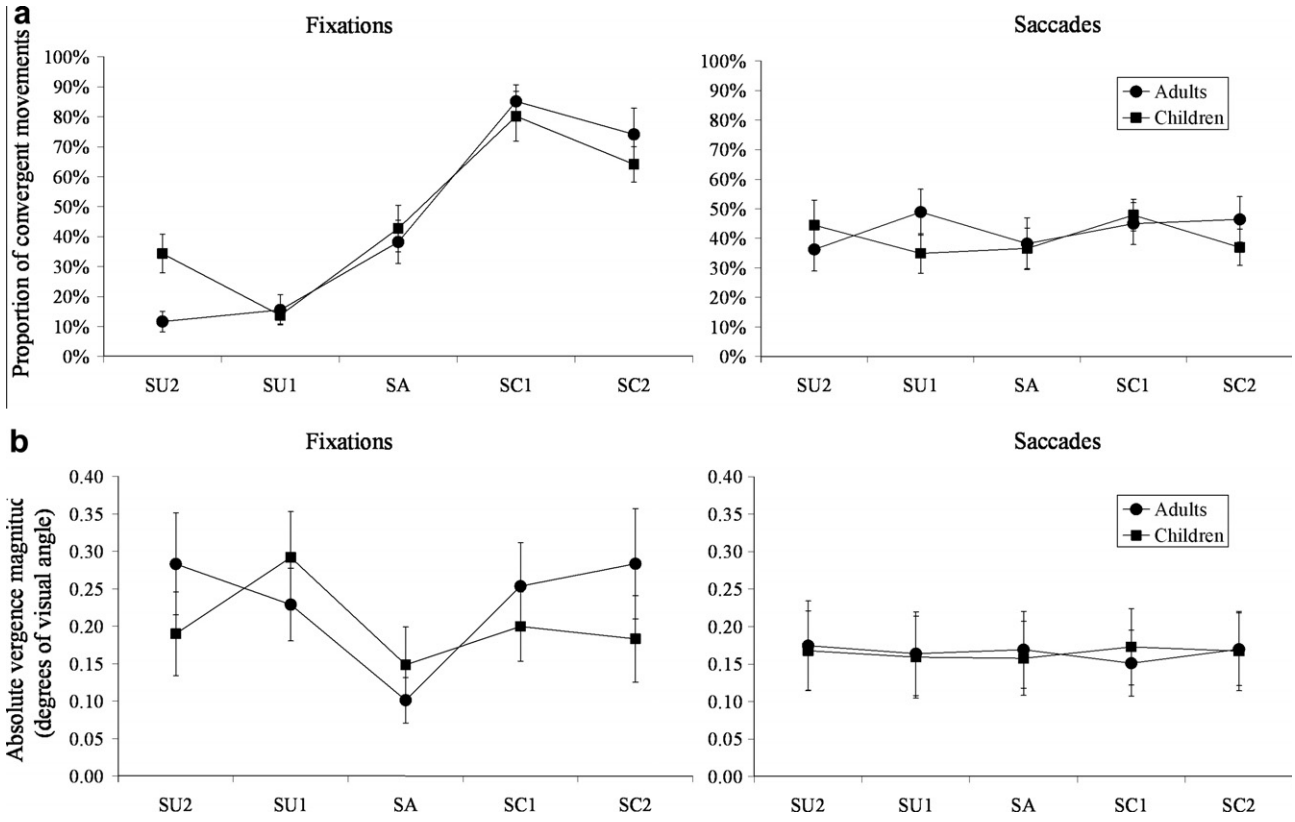

Fig. 3. A comparison of vergence movements during the initial saccade onto or during the initial fixation on the stereoscopic stimuli. The mean probability of a vergence movement being convergent (as opposed to divergent) is shown in Panel a, while the mean absolute magnitude of those vergence movements is shown in Panel b. Bars show the standard error for each participant group in each condition. On the abscissa axes (stereoscopic disparity), the letter U denotes an uncrossed disparity, the letter C denotes a crossed disparity, and A refers to an aligned stimulus. The numbers denote the magnitude of the disparity in character spaces.

reduce or eliminate retinal disparity. Thus, for adults, retinal inputs were aligned at the end of the final fixation on 38% of trials, were disparate by one character on 41% of trials (27% uncrossed, 14% crossed) and, on a minority of trials, 21%, inputs were disparate by two characters (13% uncrossed, 8% crossed). For children, retinal inputs were aligned at the end of the final fixation on 40% of trials, were disparate by one character on 33% of trials (21% uncrossed, 12% crossed), and were disparate by two characters on 27% of trials (15% uncrossed, 12% crossed). For this reason, standard errors are far greater in the TU2 and TC2 conditions compared to the other retinal disparity conditions.

These data show that, generally, when disparity is experienced, the vergence system is very efficient at reducing that disparity in support of fusion of the retinal inputs – during the majority of trials retinal disparity was reduced and, subsequently, response accuracy was high. However, when larger disparities are experienced, and the vergence system does not respond to a sufficient degree (21–27% of trials where two character spaces of retinal disparity remained at the end of the final fixation), participants are impaired in making an accurate lexical decision, presumably because the fusion process is unsuccessful.

In the remainder of the Results section, we report eye movement data. In Sections 3.2 and 3.3 we report data on binocular coordination as a function of stereoscopic disparity, in order to examine whether or not participants were sensitive to disparity cues in the parafovea compared to when they were directly fixating the stimulus. In Sections 3.4 and 3.5 we examine eye movement control as a function of total retinal disparity, our summed measure incorporating both stereoscopic disparity and fixation disparity. This allowed us to assess eye movement behaviour through the course of each trial as vergence movements were made in order to reduce retinal disparity, with the aim of fusing the stimuli.

3.2. Vergence movement in relation to saccadic targeting

We examined two aspects of vergence movements – the direction (convergent or divergent), and the magnitude of those movements (the difference in absolute fixation disparity between the start and end of the fixation or saccade) – in relation to stereoscopic disparity (the stimulus manipulation). We compared these characteristics of vergence during the initial saccade onto the stereoscopic stimulus with those during the initial fixation on the stimulus (see Fig. 3).[3] The aim was to examine whether participants were sensitive to parafoveal disparity cues and, therefore, whether saccades were targeted to words that had or had not been fused prior to direct fixation. Note that these analyses essentially amount to an assessment of whether the saccades were accurately targeted in depth (Enright, 1984, 1986, 1998). If participants were sensitive to parafoveal disparity cues and, hence, were targeting their saccades to a stimulus in apparent depth, then the data ought to show some systematic effect of stereoscopic disparity on the vergence change during the initial saccade onto the stimulus.

There were no overall effects of group or of vergence type (initial saccade vs. initial fixation) (both Fs < 1). There was, however, a significant effect of stereoscopic disparity as well as a significant interaction between stereoscopic disparity and vergence type (both Fs > 25, both ps < .001). Very clearly, participants made stimulus-appropriate vergence movements following the point at which the target was directly fixated, whereas there was little, if any, evidence of such vergence during the saccade onto the target. This pattern of effects held for both adults and children. Once the

[3] In this analysis, we only included data from the first fixation on the letter string from each trial. More comprehensive analyses of vergence based on every fixation made on the letter strings are reported in Section 3.5.

participant was directly fixating the stereoscopic stimulus the vergence system responded strongly to disparities; vergence movements were both direction- and magnitude-appropriate (all $ts > 2$, all $ps \leqslant .01$). In contrast, the binocular coordination of the eyes was not in any way modulated by the disparity characteristics of the target during the saccade, clearly demonstrating that participants only made vergence movements to fuse the stimuli once they were directly fixating the stimulus. To reiterate, there was no evidence that the disparity characteristics of non-foveal stimuli influenced binocular saccade metrics.

Recall that if vergence during the initial saccade was driven by the apparent depth of the stimulus then this would suggest participants had fused the stereoscopic stimulus in the parafovea (in order to appropriately target their saccade in depth). In contrast, however, the lack of an effect of stereoscopic disparity on initial saccade vergence suggests that participants were not able to fuse the disparate stimuli prior to direct fixation; rather, they were targeting their saccades to a non-fused target string.

3.3. Landing positions

In relation to our analyses of vergence movements during the initial saccade onto the target, we also examined landing positions on the target. Here, rather than examining the relative change in the two eyes' alignment from the start to the end of the initial saccade, we examined where the two eyes landed within the target letter string in relation to each other – the absolute positions of the eyes which resulted from a change in vergence during the incoming saccade, and from which subsequent vergence movements were made during the initial fixation on the word. These analyses are highly complementary to those reported in Section 3.2 and were based on a similar prediction; if participants were sensitive to parafoveal disparity cues then landing positions of the two eyes (as measured at the depth of the screen) ought to be systematically modulated by the stereoscopic disparity of the stimulus; (a) the greater the magnitude of disparity in the stereoscopic stimuli, the greater the difference in the landing positions of the two eyes; (b) the direction of disparity in the two eyes' landing positions (crossed or uncrossed) ought to be determined by the direction of stereoscopic disparity.

The data are shown in Fig. 4. (Recall that one character space subtended 0.37°.) Thus, the eyes were, on average, landing between the second and third characters of the stimulus (mean landing position, averaged between the positions of the two eyes, was 2.6 characters for the adults, 2.5 characters for the children).

As can clearly be seen in Fig. 4, landing positions for the two eyes were not the same – the right eye landed consistently further to the right within the target than the left eye ($F(1, 18) = 10.49$, $p = .01$). This finding is entirely consistent with previous work showing that the two eyes are predominantly uncrossed at the start of a fixation (Blythe et al., 2006; Liversedge, Rayner, et al., 2006; Liversedge, White, et al., 2006). While there was no overall difference between adults and children ($F(1, 18) < 1$), there was a main effect of stereoscopic disparity on landing positions ($F(4, 72) = 20.61$, $p < .001$). As can be seen in Fig. 4, the greater the magnitude of stereoscopic disparity the further from the left edge of the stereoscopic letter string the eyes landed. Given the lack of an effect of stereoscopic disparity on binocular coordination (the difference in the two eyes' landing positions did not change across the different stereoscopic disparity conditions, $p > .1$), the main effect of stereoscopic disparity was not related to differences in the two eyes' retinal inputs.

Rather, it seems likely that this effect is related to the apparent size of the stimulus in the parafovea. All the stimuli were six letters long; with one character space of stereoscopic disparity the non-fused target would have taken up seven character spaces on the display screen, and with two character spaces of stereoscopic disparity the non-fused target would have taken up eight character spaces on the display screen. The greater the magnitude of stereoscopic disparity, the fewer letters would have overlapped between the two images and, hence, the longer the parafoveal target would have appeared. Recall that there was no effect of parafoveal stereoscopic disparity on the change in vergence during the initial saccade onto the stimulus (Section 3.2). This indicates that participants were targeting their saccades to a non-fused target letter string. Thus, the change in landing positions reflects the oculomotor system targeting the preferred viewing position (O'Regan, 1981; Rayner, 1979) of a string that was not fused prior to direct fixation and, therefore, varied in length systematically in relation to the magnitude of the stereoscopic manipulation.

3.4. The time course of fusion

Recall that once the target word had appeared it remained on the screen until the participant pressed a button; hence, it was quite common for multiple fixations to be made on the word within any one trial. The participants reported that during each trial the stimulus initially appeared diplopic and then after a few moments became clear. In relation to this subjective experience, the multiple fixations within a trial correspond to the period over which partic-

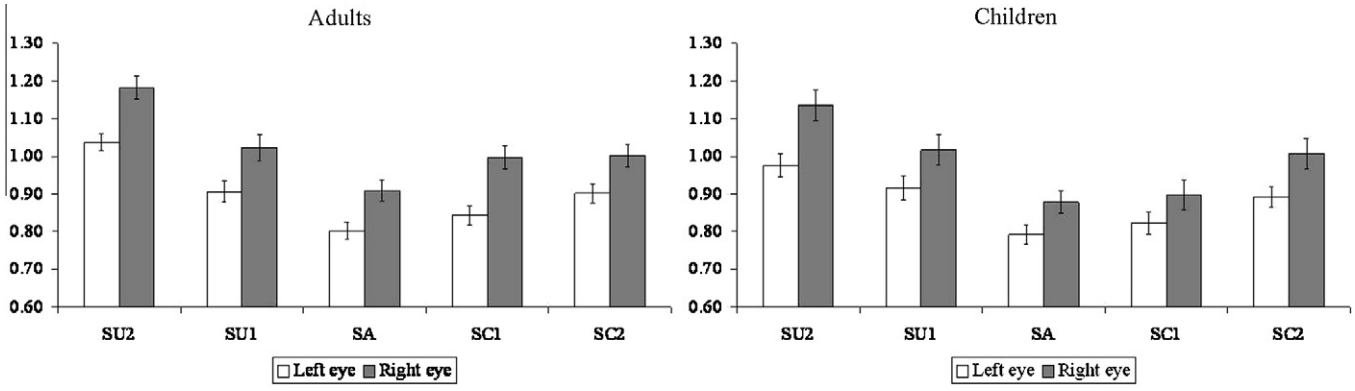

Fig. 4. Mean landing positions of the two eyes on the target letter string following the initial saccade from the fixation cross. The ordinate axis shows the mean landing position in degrees of visual angle; recall that one character space subtended .37°. On the abscissa axes (stereoscopic disparity), the letter U denotes an uncrossed disparity, the letter C denotes a crossed disparity, and A refers to an aligned stimulus. The numbers denote the magnitude of the disparity in character spaces. Bars show the standard error for each participant group in each condition.

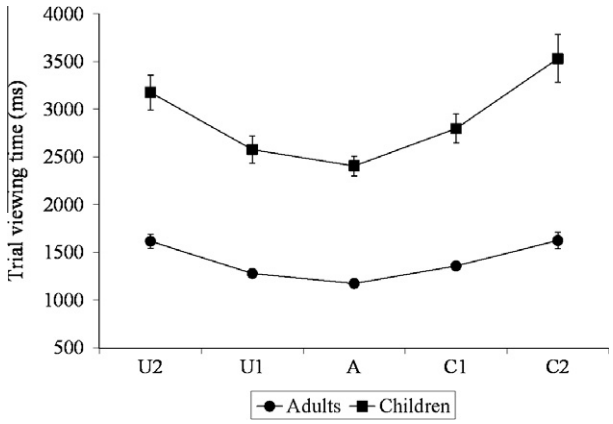

Fig. 5. Total trial viewing times for adults and children as a function of total retinal disparity at the start of the first fixation of each trial. On the abscissa axes (total retinal disparity), the letter U denotes an uncrossed disparity, the letter C denotes a crossed disparity, and A refers to an aligned stimulus. The numbers denote the magnitude of the disparity in character spaces. Bars show the standard error for each participant group in each condition.

ipants made vergence movements in order to fuse the diplopic stimulus.

Total trial viewing time and the number of fixations per trial were measured as a function of retinal disparity during the first fixation of each trial (shown in Fig. 5 and Table 1). By examining retinal disparity from the first fixation of the trial, we have an index of the direction and magnitude of vergence movements that occur at the onset of direct fixation of the stimulus. In contrast to disparity at the end of the final fixation, this initial disparity for each trial does not describe the effective fusional range since fusion may not yet have been achieved. Both total trial viewing time and the number of fixations per trial were calculated from the moment that the letter string appeared until the moment that the participant made their lexical decision.

It is immediately striking that trial viewing times are quite long in comparison to standard lexical decision tasks (e.g. 500–700 ms, Schilling, Rayner, & Chumbley, 1998), and there are several reasons for this. First, these response times include the time to initiate the initial saccade from the location of the fixation cross to the target letter string and the duration of the saccade itself; more typically in lexical decision tasks, participants are not required to saccade onto the target location at the beginning of each trial. Second, in four out of five trials the stimulus was presented with stereoscopic disparity and initially appeared diplopic; thus, the increased decision times are also related to the time taken for participants to make vergence movements and to fuse the stimulus. Third, the task instructions differed slightly from a standard lexical decision task in that participants were required to decide whether the word was correctly spelled. The impact of this task instruction is clearest in the case of aligned retinal inputs, where minimal vergence movements were required, and yet decision times were still relatively long – 1175 ms for the adults and 2406 ms for the children. Note that our task was selected with the intention that detecting single-letter errors would make participants more likely to wait until they had fully fused the stimulus before making their decision (and so, less likely to make guesses). It seems plausible, therefore, that searching for spelling errors would lead to longer decision times than simply deciding whether a letter string is a real word or not.

Unsurprisingly, the patterns of data for the total viewing time and the number of fixations per trial were very similar (since these two measures are highly correlated). Participants made more fixations per trial, and hence had longer trial viewing times, when retinal disparity was greater (Fs > 8, ps < .001). This effect clearly demonstrates increased difficulty associated with fusing larger retinal disparities. Decision times for aligned inputs were shortest, with those for disparities of one character space being somewhat longer, and those for disparities of two character spaces longest. Disparities of two character spaces in either direction led to a significant increase in the number of fixations and the total viewing time per trial compared to the other three conditions (ts > 2, ps < .01). Again, these data reflect increased difficulty in fusing stimuli with greater magnitudes of retinal disparity. One character space of retinal disparity was associated with a mean trial viewing time increase of 143 ms in adults and 283 ms in children, and two character spaces of retinal disparity was associated with an increase of 446 ms in adults and 949 ms in children.

Children made more fixations and had longer total viewing times than adults (Fs > 16, ps $\leqslant$ .001), indicating that overall they found the task more difficult than the adults. Although the interaction between participant group and retinal disparity for number of fixations was not reliable ($F(4, 72) = 1.67$, $p = .17$), the patterns of effects were very similar to those obtained for total viewing times for which the interaction was reliable ($F(4, 72) = 2.82$, $p = .03$). While the pattern of effects was similar in nature for adults and children on both measures, it is clear that the magnitude of the effect between disparities of one and two character spaces was more pronounced in children than in adults. Note that this was the case even though the children's data were noisier than those of the adults.

Hence, for total viewing times, due to increased standard errors, the difference between aligned inputs and disparities of one character space in either direction was not significant for children (ts < 2, ps > .1), while a slightly smaller numerical difference was statistically reliable for the adults (ts > 4, ps < .01). The comparison of two-character disparities with aligned inputs was significant for both participant groups (ts > 3, ps $\leqslant$ .01), though note that this effect was 446 ms for the adults compared to an effect of 949 ms in the children.

With respect to the time course of fusion, it is clear from these data that increased retinal disparity led to participants making more fixations and having longer trial viewing times. This was particularly pronounced for disparities of two character spaces in either direction compared to the aligned case. Finally, the increased difficulty associated with retinal disparities of two character spaces was greater in children compared to adults. These data correspond to the period during which participants reported that they experienced an initially diplopic stimulus that became non-

Table 1
Mean total trial viewing time (ms) and mean number of fixations per trial across different values of retinal disparity. The letter U denotes an uncrossed disparity; the letter C denotes a crossed disparity; A refers to aligned retinal inputs. The numbers denote the magnitude of the disparity in character spaces. Standard errors are shown in parentheses.

		Total retinal disparity				
		U2	U1	A	C1	C2
Total trial viewing time	Adults	1617 (72)	1278 (33)	1175 (30)	1358 (43)	1624 (87)
	Children	3178 (183)	2578 (144)	2406 (104)	2799 (152)	3531 (215)
Number of fixations per trial	Adults	3.8 (.1)	3.3 (.1)	3.2 (.1)	3.3 (.1)	3.8 (.2)
	Children	5.4 (.3)	4.5 (.2)	4.6 (.2)	5.4 (.3)	6.0 (.4)

diplopic prior to them making a response. It seems likely that these longer trial viewing times and numbers of fixations were associated with the need for more vergence movements, and those movements to be of greater magnitude, in order to reduce retinal disparity, allowing participants to accurately complete the task.

3.5. Vergence movements during fixations

We report here two measures – the probability of vergence movements being in the appropriate direction to reduce retinal disparity, and the magnitude of those vergence movements. The impact of retinal disparity on the time course of fusion seems likely to be related to the time necessary for participants to make appropriate vergence movements during fixations. We predicted, therefore, that more direction-appropriate vergence movements would be made, and those movements would be of greater magnitude, when participants were experiencing increased retinal disparity.

Note that the occurrence of vergence movements cannot necessarily be taken as an indication that the stereoscopic stimulus was unfused prior to the initiation of such vergence movements, as vergence can be elicited by disparities within Panum's fusional area (Masson, Busettini, & Miles, 1997; Riggs & Niehl, 1960). Furthermore, research examining binocular coordination during reading has found that vergence movements typically occur during fixations despite the fact that the stimulus is presented non-stereoscopically, and fixation disparity does not cause the words to appear diplopic (Blythe et al., 2006; Kirkby et al., 2008; Liversedge, Rayner, et al., 2006; Liversedge, White, et al., 2006). However, given the magnitude of disparity that was manipulated in the present study, along with participants' reports that the stimuli initially appeared unfused, and the inappropriate binocular alignment during the initial saccade as well as landing positions on the stimuli, it seems likely that the vergence movements described here were being elicited by an unfused stimulus.

In support of this suggestion, and for illustrative purposes, the raw data from one adult participant on three different trials are shown in Fig. 6. Panel a shows data from a trial where the participant was viewing a letter string with 0 stereoscopic disparity. The two eyes are slightly uncrossed, and this remains relatively constant throughout the trial. Panel b shows data from a trial where the stimulus was presented with two character spaces of uncrossed disparity. Here, the two eyes are initially slightly uncrossed and during the course of the trial eyes diverge such that the magnitude of uncrossed disparity increases. Finally, Panel c shows data from a trial where the stimulus was presented with two character spaces of crossed disparity. Again, the eyes are initially slightly uncrossed but they converge during the trial such that, by the time the participant made their lexical decision, the two eyes are crossed (though remaining slightly uncrossed relative to stereoscopic disparity).

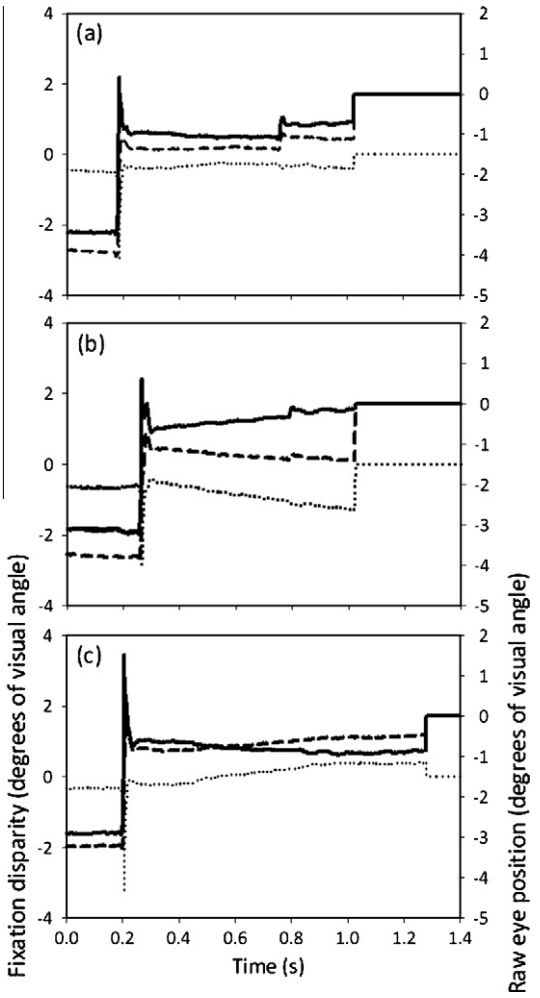

Fig. 6. The raw eye positions, and corresponding vergence angle, of one adult participant for three trials. Panel a shows a trial where the stereoscopic disparity was aligned, Panel b shows a trial where the stereoscopic disparity was uncrossed by two character spaces and the Panel c shows a trial where the stereoscopic stimulus was crossed by two character spaces. The fine dotted lines represent fixation disparity between the two eyes, the dashed lines represent the position of the left eye and the solid lines represent the position of the right eye. The ordinate axis on the left side corresponds to fixation disparity, with positive values denoting crossed fixations and negative values denoting uncrossed fixations. The ordinate axis on the right side corresponds to the eye positions where a value of 0 shows the centre of the screen, negative values denote positions on the left side of the screen, and positive values denote positions on the right side of the screen. The overshoot shown following the first saccade is accountable for by the lens slip artefact found when using the dual Purkinje recording method (Deubel & Bridgeman, 1995).

First, we compared the probability of movements being convergent (as opposed to divergent) in relation to retinal disparity. Unsurprisingly, the direction of vergence movements was corrective for the direction of retinal disparity ($F(4, 72) = 52.76$, $p < .001$). When retinal disparity was uncrossed (i.e. stereoscopic disparity was crossed relative to the vergence state), convergence was most likely, and when the disparity was crossed (i.e. stereoscopic disparity was uncrossed relative to the vergence state), divergence was most likely (see Panel b of Fig. 3). Although there was no difference in overall direction-appropriate vergence probability between adults and children ($F < 1$), we found a significant interaction between group and disparity ($F(4, 72) = 8.60$, $p < .001$) for this measure. Adults made a higher proportion of direction-appropriate vergence movements compared to the aligned condition for both crossed and uncrossed disparities of one and two characters ($ts > 2$, $ps \leqslant .02$). However, while children made a numerically higher proportion of direction-appropriate vergence movements for crossed and uncrossed disparities of one character space ($ts > 2$, $ps \leqslant .01$), vergence movements were not significantly different to those observed under aligned conditions for crossed and uncrossed disparities of two characters ($ts < 2$, $ps > .1$).

Thus, when experiencing one character space of retinal disparity, all participants made convergent movements to reduced uncrossed disparity (i.e. when the stereoscopic disparity was crossed relative to the actual vergence state) and divergent movements to reduce crossed disparity (i.e. when the stereoscopic disparity was uncrossed relative to the actual vergence state). These data show that the direction of the retinal disparity determined the direction of vergence movements, such that both adults and children made corrective movements to reduce it. This was also the case for the adult participants when experiencing two character spaces of retinal disparity. However, when children experi-

enced two character spaces of disparity, the probability of their vergence movements being direction-appropriate was no different to that which occurred in the aligned condition.

Next, we computed the mean absolute magnitude of vergence movements in relation to retinal disparity. This analysis allowed us to examine whether participants were making larger vergence movements in response to greater magnitudes of retinal disparity (again, see Panel b of Fig. 3). We found that the magnitude of vergence movements increased in response to larger retinal disparities ($F(4, 72) = 11.35$, $p < .001$). For adults, the mean vergence magnitude was 0.25°, 0.16°, 0.12°, 0.21°, and 0.27° for disparities of TC2, TC1, A, U1, and U2 character spaces, respectively. For children, the mean vergence magnitude was 0.18°, 0.17°, 0.12°, 0.21°, and 0.15° for disparities of TC2, TC1, A, U1, and U2 character spaces, respectively. The adult data clearly show that the magnitude of vergence movements was driven by the magnitude of retinal disparity, such that larger vergence movements were made in response to increased retinal disparity. The children's data show increased magnitudes of vergence movements in response to one character space retinal disparities, but there was no corresponding increase in vergence magnitude for retinal disparities of two character spaces.

While there was a marginal overall difference in the magnitude of vergence movements between adults and children ($F(1, 18) = 4.18$, $p = .06$), most importantly, there was a highly reliable interaction between group and disparity ($F(4, 72) = 5.03$, $p = .001$). The pattern of effects was extremely similar to that observed for the proportion of direction-appropriate vergence movements. In adults, the magnitude of vergence movements was least for aligned retinal inputs, and increased proportionally with increasing degrees of crossed and uncrossed disparity (all $ts \geqslant 3$, all $ps \leqslant .01$). For one character space of disparity vergence magnitude was, on average, .19°, while for two character spaces of disparity vergence magnitude was, on average, 0.26°.

For the children, the aligned and one character crossed and uncrossed data patterned almost identically to those for the adults (both $ts > 4$, both $ps < .01$). In contrast, however, children were much less efficient than adults in making vergence movements in response to both crossed and uncrossed disparities of two characters. The difference in the magnitude of vergence movements for two character uncrossed or crossed disparities was reliably different from that observed for aligned inputs ($ts > 2$, $ps < 0.05$). However, there was clearly no proportional increase in vergence magnitude between the one and two character space conditions as was the case for the adults. For the children, vergence magnitude in response to one character space of disparity was, on average, .19°, and was actually slightly smaller (.17°) in response to two character spaces of disparity.

Thus, while adults made vergence movements that were proportional to the degree of disparity they experienced for the full range of disparities tested here, the magnitude of vergence movements in children did not appropriately increase for disparities of two characters (crossed or uncrossed). Again, it appears that children were less responsive in terms of vergence eye movements for disparities of greater magnitude when compared with adults.

These data can be considered in relation to those from the analysis of the time course of fusion (Section 3.2). Adult participants made effective vergence movements in response to retinal disparities of both one and two character spaces. They also make more fixations, and the time course of fusion was longer for larger disparities. In contrast, children were less efficient in making vergence movements in response to larger retinal disparities in terms of both direction and magnitude. In sum, children took longer to make a decision, and their vergence response was reduced such that it was less effective, at greater compared with smaller disparities.

4. Discussion

We can summarise the events that occurred in our experiment when the participants made a lexical decision about a word presented stereoscopically in the visual periphery as follows. An orienting saccade was made from a cross on the left of the display screen to bring the word into foveal vision. When the word was presented with stereoscopic disparity, vergence appropriate to reduce this disparity occurred from the onset of the following fixation. We found no evidence that vergence movements were initiated before or during the saccade. Adults made an appropriate vergence response to both large (two character spaces) and small (one character space) retinal disparity, whereas children responded appropriately for small disparities, though less well to large disparities. In cases where disparity was reduced by vergence movements to one character space or less by the end of the final fixation, an accurate response was usually made in the lexical decision task. However, when retinal disparity remained at more than one character space, response accuracy was significantly impaired in adults, and even more so children.

A number of interesting conclusions can be drawn on the basis of these data. Both adults and children were able to make reliable lexical decisions on stereoscopically presented letter strings. While there is a large experimental literature based on the lexical decision task, and several well-developed computational models of lexical identification, this study is, to our knowledge, the first report of lexical decision data for stereoscopically presented stimuli. The data show that, despite inducing longer response times (Section 3.4), the stereoscopic presentation in itself did not prevent participants from making reliable lexical decisions.

We investigated the effective fusional range by examining retinal disparities over which participants were consistently and reliably able to fuse stereoscopically presented words. For both adults and children, the data clearly demonstrated that participants were able to successfully fuse linguistic stimuli with up to 0.37° (one character space) of crossed or uncrossed disparity. Within this range of disparities, response accuracy was extremely high – within 5% of accuracy in the case of perfectly aligned retinal inputs (for which performance was optimal). In contrast, when two character space (0.74°) retinal disparities remained at the end of the final fixation, response accuracy was significantly impaired. For this reason, we estimate the effective fusional range for linguistic stimuli presented in a natural viewing situation to be approximately 0.37° and, importantly, the same limits applies to both adults and children. Contrary to our hypotheses, we did not find that children were able to fuse a greater range of disparity magnitudes than adults. However, this finding of an equal fusional range in children and adults relates to the following conclusions concerning their vergence responses to retinal disparity.

The adults made direction- and magnitude-appropriate vergence movements, reducing the experienced disparity. These vergence movements were effective in reducing disparities of both one and two character spaces through the course of the trial, such that relatively little disparity remained at the moment when they made their lexical decision. Thus, although 40% of the stimuli were presented with two character spaces of stereoscopic disparity, retinal disparities of two character spaces remained by the end of the final fixation on just 21% of trials for adults. These data can be clearly related to the response accuracy data. In those cases where they had failed to make such vergence movements (i.e. where there remained a disparity of two character spaces by the end of the trial), response accuracy dropped to 86%.

Children also made direction- and magnitude-appropriate vergence responses to retinal disparities of one character space. In contrast, though, they made fewer direction-appropriate vergence movements, and the magnitude of vergence movements in

response to two character space disparities was actually smaller than that which occurred for disparities of one character space. Thus, in comparison with adults, children's vergence responses to large retinal disparities were significantly less effective, in terms of both direction and magnitude, and as a consequence they much more frequently failed to reduce the experienced disparity through the course of the trial.

There are three important points to note with respect to children's vergence response to large retinal disparities, though. First, similar to the adult data, there is a clear link between the vergence response and subsequent response accuracy (indexing the success of fusion). Two character spaces of retinal disparity remained at the end of the final fixation on 27% of trials for children and, in these cases where vergence movements had failed to reduce retinal disparity, response accuracy was just 65%. Second, these data are consistent with the literature showing that binocular coordination is poorer in children than in adults on both reading and non-reading tasks (Blythe et al., 2006; Fioravanti et al., 1995; Yang & Kapoula, 2003; Yang et al., 2002). Third, the average disparity that occurs during fixations in normal reading for children is .3° (Blythe et al., 2006) – less than the one character space disparity manipulations in the present study (recall that one character space = 0.37°). Thus, the data from this study provide compelling evidence that, despite their relatively poor binocular coordination in comparison with adults, children are still able to fuse the range of fixation disparities that they typically experience during normal reading.

To summarise, the relative decrement in response accuracy from one to two character spaces of disparity was far greater for children than for adults, and this would appear to be a direct consequence of children's failure to make appropriate vergence movements in response to large disparities. Response accuracy for both adults and children was near-optimal for disparities of up to one character space. However, children's response accuracy, relative to that of adults, decreased substantially for disparities of two character spaces (a decrease of 21% in children compared to 9% in adults), and this was as a consequence of their failure to make appropriate vergence movements. We have shown both consistencies and differences between adults and children in terms of vergence responses to disparities. For children, the vergence system is sensitive and responsive to disparities that can be fused. However, for disparities beyond the effective fusional range, the vergence system is far less effective. Adults are similarly sensitive and responsive to disparities that can be fused; however, in contrast with children, appropriate vergence responses in adults are also elicited by disparities that cannot easily be fused.

The final conclusion relates to our examination of parafoveal and foveal processing of stereoscopic disparity. We found, very clearly, that participants did not use parafoveal disparity cues to target their saccades – there was no evidence for an effect of parafoveal disparity either in the analysis of vergence changes during the initial saccade, or in the analysis of initial landing positions on the stereoscopic letter strings. In contrast, the vergence response was clearly driven by retinal disparity once the participants were directly fixating the stimulus. Thus, binocular coordination was modulated by foveal but not by parafoveal disparity cues. A related point is that participants were able to target their saccades to the preferred viewing position – just left of the centre of the letter string – for stereoscopically presented letter strings. Again, to our knowledge, this is the first demonstration of what might be considered "normal" oculomotor behaviour in relation to linguistic stimuli, using a stereoscopic presentation technique.

There are several key benefits to the methodological approach that we adopted in this study. First, by examining response accuracy in relation to eye movement data, we have an objective, behavioural measure of the process of fusion where previous studies have relied on subjective self-reports. We felt that the use of self-report with young children, in particular, could have been unreliable.

Second, it allows us to eliminate the possibility that non-diplopic vision is being achieved through a process of suppression rather than fusion. Were it the case that participants were suppressing one of the two retinal inputs then there would be no influence of the stereoscopic disparity manipulation on the eye movement data; this manipulation would only serve to move the position of the stimulus to the non-suppressed eye slightly to the left or right from one trial to the next. Furthermore, were non-diplopic vision were being attained through suppression then the stereoscopic manipulation would not have affected the ease with which participants were able to perform the lexical decision task, and so there would not have been any effect on either the response accuracy or eye movement measures. This was not the case. As retinal disparity increased, response accuracy and eye movement data both showed substantial effects of retinal disparity; these effects can only be explained by an underlying process of fusion and not suppression.

An important point to consider is whether participants might be able to successfully complete the task without fusing the stereoscopic stimuli. Based on the subjective experience of the authors when viewing these stimuli through the shutter goggles, it was simply not possible to distinguish the individual letters of the stimulus when first looking at the letter strings (hence, any guessing strategy would have resulted in a response accuracy close to 50%). The words appeared as a diplopic image, with at least four of the six letters of the word overlapping. After looking at the target for a few moments, the stimulus quite suddenly appeared clearly as a single word and it was then extremely simple to make the lexical decision. These subjective reports are supported by the eye movement data (see examples in Fig. 6) in which large vergence movements were made during the trial which aligned the eyes roughly onto the stimulus, allowing fusion to occur. Given that the nonword stimuli were created by substituting a single letter in the centre of each word (i.e. always within the overlapping portion of the stereoscopic stimulus), it seems implausible that lexical decisions could be reliably made on the basis of any guessing strategy, were participants failing to fuse the stimuli. None of the participants reported to us that they had been guessing; all comments in feedback indicated that they were generally able to see the words quite clearly.

Finally, as can be seen in Panel c of Fig. 1, when the stimulus was presented with two character spaces of stereoscopic disparity then the participant was able to see the first and the last two letters of the six letter word quite clearly without fusing the stimulus. Importantly, the overlapping portion of the word was four characters long. In contrast, when the stimulus was presented with one character space of stereoscopic disparity then the participant was able to see only the first and the sixth letter clearly without fusing, and the overlapping portion of the word in these trials was five characters long. Therefore, any attempt to complete the task without successfully fusing the stimulus ought to have lead to better response accuracy in the TU2 and TC2 conditions than in the TU1 and TC1 conditions; this was clearly not the case.

In summary, these data represent the first direct measurement of binocular vergence responses to stereoscopically disparate stimuli in children and adults under directly comparable experimental conditions. Our results have allowed us to investigate for the first time the effective fusional range for both adults and children in relation to visual linguistic stimuli presented in a natural viewing situation. Our characterisation of the vergence response to stereoscopically presented letter strings is similarly a new finding, and is of particular importance given developmental changes that occur between beginning and skilled readers in the complex neurophysiological system that is responsible for binocular coordination

(Blythe et al., 2006; Yang & Kapoula, 2003; Yang et al., 2002). Very fundamentally, the present study represents a demonstration of the centrality of vergence responses to the formation of a single unified perceptual representation of the visual environment within the human visual system.

Acknowledgments

This research is based on a doctoral dissertation by Hazel I. Blythe at the University of Southampton, initiated at Durham University. Portions of the data were presented at the 2007 meeting of the Psychonomic Society in Long Beach, and appear in the Proceedings of the 2009 SPIE Stereoscopic Displays and Applications conference in San Jose. This research was supported by Grants F/00 128/AB and F/00 128/AG from the Leverhulme Trust, and Grant G0601757 from the MRC to S. Liversedge. H. Blythe was supported by an ESRC Postdoctoral Fellowship. We would like to thank Bob Metcalf for programming assistance, and the parents and children who took part in this experiment as unpaid volunteers. We would also like to thank Keith Rayner and Wolfgang Jaschinski for their helpful comments on this paper.

References

Ames, A., Ogle, K. N., & Gliddon, G. H. (1932). Corresponding retinal points, the horopter and size and shape of ocular images. *Journal of the Optical Society of America, 22*, 575–631.

Bains, R. A., Crawford, J. D., Cadera, W., & Vilis, T. (1992). The conjugacy of human saccadic eye movements. *Vision Research, 32*, 1677–1684.

Barlow, H., Blakemore, C., & Pettigrew, J. (1967). The neural mechanisms of binocular depth discrimination. *Journal of Physiology, 193*, 327–342.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vision Research, 46*, 3898–3908.

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science, 208*, 615–617.

Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988). Binocular coordination of human horizontal saccadic eye movements. *Journal of Physiology, 404*, 157–182.

Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1995). Voluntary binocular gaze-shifts in the plane of regard: Dynamics of version and vergence. *Vision Research, 35*, 3335–3358.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

Cova, A., & Galiana, H. L. (1995). Providing distinct vergence and version dynamics in a bilateral oculomotor network. *Vision Research, 35*, 3359–3371.

DeAnglis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in Macaque area MT. *The Journal of Neuroscience, 19*, 1398–1415.

Deubel, H., & Bridgeman, B. (1995). Fourth Purkinje image signals reveal lens deviations and retinal image distortions during saccadic eye movements. *Vision Research, 35*, 529–538.

Dowd, J. M., Clifton, R. K., Anderson, D. R., & Eichelman, W. H. (1980). Children perceive large-disparity random-dot stereograms more readily than adults. *Journal of Experimental Child Psychology, 29*, 1–11.

Enright, J. T. (1984). Changes in vergence mediated by saccades. *Journal of Physiology, 350*, 9–31.

Enright, J. T. (1986). Facilitation of vergence changes by saccades: Influences of misfocused images and of disparity stimuli in man. *Journal of Physiology, 371*, 69–87.

Enright, J. T. (1998). Monocularly programmed human saccades during vergence changes? *Journal of Physiology, 512*, 235–250.

Erkelens, C. J., & Sloot, O. B. (1995). Initial directions and landing positions of binocular saccades. *Vision Research, 35*, 3297–3303.

Fender, D., & Julesz, B. (1967). Extension of Panum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America, 57*, 819–830.

Fioravanti, F., Inchingolo, P., Pensiero, S., & Spanios, M. (1995). Saccadic eye movement conjugation in children. *Vision Research, 35*, 3217–3228.

Heckmann, T., & Schor, C. M. (1989). Panum's fusional area estimated with a criterion-free technique. *Perception & Psychophysics, 45*, 297–306.

Howard, I. P. (2001). *Seeing in depth, basic mechanisms* (Vol. 1). Canada: Porteous.

Judge, S. J., & Cumming, B. G. (1986). Neurons in the monkey midbrain with activity related to vergence eye movements and accommodation. *Journal of Neurophysiology, 55*, 915–930.

Kenyon, R. V., Ciuffreda, K. J., & Stark, L. (1980). Unequal saccades during vergence. *American Journal of Optometry and Physiological Optics, 57*, 586–594.

King, W. M., & Zhou, W. (2000). New ideas about the binocular coordination of eye movements: Is there a chameleon in the primate family tree? *The Anatomical Record, 261*, 153–161.

Kirkby, J. A., Webster, L. A. D., Blythe, H. I., & Liversedge, S. P. (2008). Binocular coordination during reading and non-reading tasks. *Psychological Bulletin, 134*, 742–763.

Kloke, W. B., & Jaschinski, W. (2006). Individual differences in the asymmetry of binocular saccades, analysed with mixed-effects models. *Biological Psychology, 73*, 220–226.

Lee, H. S., & Dobbins, A. C. (2006). Perceiving surfaces in depth beyond the fusion limits of their elements. *Perception, 35*, 31–39.

Leigh, R. J., & Zee, D. S. (1999). *The neurology of eye movements*. Oxford, UK: Oxford University Press.

Liversedge, S. P., Rayner, K., White, S. J., Findlay, J. M., & McSorley, E. (2006). Binocular coordination of the eyes during reading. *Current Biology, 16*, 1726–1729.

Liversedge, S. P., White, S. J., Findlay, J. M., & Rayner, K. (2006). Binocular coordination of eye movements during reading. *Vision Research, 46*, 2363–2374.

Masson, G. S., Busettini, C., & Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature, 389*, 283–286.

Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9*, 1–28.

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44*, 253–259.

O'Regan, J. K. (1981). The "convenient viewing position" hypothesis. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 289–298). Hillsdale, NJ: Erlbaum.

Palmer, D. A. (1961). Measurement of the horizontal extent of Panum's Area by a method of constant stimuli. *Journal of Modern Optics, 8*, 151–159.

Pettigrew, J. D., Nikara, T., & Bishop, P. O. (1968). Binocular interaction on single units in cat striate cortex: Simultaneous stimulation by moving slit with receptive fields in correspondence. *Experimental Brain Research, 6*, 353–372.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*, 21–30.

Richards, W., & Foley, J. M. (1971). Interhemispheric processing of binocular disparity. *Journal of the Optical Society of America, 61*, 419–421.

Riggs, L. A., & Niehl, E. W. (1960). Eye movements recorded during convergence and divergence. *Journal of the Optical Society of America, 50*, 913–920.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26*, 1270–1281.

Schor, C. M., & Tyler, C. W. (1981). Spatio-temporal properties of Panum's fusional area. *Vision Research, 21*, 683–692.

Schor, C. M., Wood, I. C., & Ogawa, J. (1984). Binocular sensory fusion is limited by spatial resolution. *Vision Research, 24*, 661–665.

Sylvestre, P. A., Choi, J. T. L., & Cullen, K. E. (2003). Discharge dynamics of oculomotor neural integrator neurons during conjugate and disjunctive saccades and fixation. *Journal of Neurophysiology, 90*, 739–754.

Yang, Q., Bucci, M. P., & Kapoula, Z. (2002). The latency of saccades, vergence, and combined eye movements in children and in adults. *Investigative Ophthalmology, 43*, 2939–2949.

Yang, Q., & Kapoula, Z. (2003). Binocular coordination of saccades at far and near in children and in adults. *Journal of Vision, 3*, 554–561.

Zee, D. S., Fitzgibbon, E. J., & Optican, L. M. (1992). Saccade–vergence interactions in humans. *Journal of Neurophysiology, 68*, 1624–1641.